Detection and Counting of Lentil Grains using Convex Deficiency for better Quality Estimation

Yumnam Kirani Singh[1], Amitava Akuli[2]

[1]Centre for Development of Advanced Computing,
Plot-E2/1, Block GP, Sector-V, Salt lake Electronics complex, Kolkata -700091
yumnam.singh@cdac.in

[2]Centre for Development of Advanced Computing
Plot-E2/1, Block GP, Sector-V, Salt lake Electronics complex, Kolkata -700091
amitava.akuli@cdac.in

***Abstract*:**
Proposed here is a simple and effective method for detection and counting of connected grains in an image. The grains considered here are the whole lentil grains which are circular in shapes. These circular shapes can be considered as convex polygons. When two or more such convex polygons are connected or partially overlapped, there are concave regions along the boundary lines of the connected region. Such concave regions form the convex deficiencies for the convex hull of the connected or overlapped grains. Finding the number convex deficiencies in the convex hull of connected object boundaries allows us to identify whether connected component is formed by connecting two or more objects or grains. Also, it allows us to exactly count the number of connected objects and simplifies the finding the segmentation boundaries. The method is simple, fast and gives good accuracy of counting connected grains without actually segmenting the connected grains. The method can successfully count the lentil grains in an image which will be useful for counting of any convex shape objects in an image.

***Keywords*: Convex hull, Convex Deficiency, YCbCr color model, Image segmentation, Connected components, Counting of lentil grains, Quality estimation.**

I. Introduction

In the recent years, the quality estimation of the food grains based on image processing has been one of the popular research areas. Several devices have also been developed for quality estimation of food grains, fruits and vegetables etc. Of the food grains lentil is one of the most popularly consumed grains in India. The quality of the lentil grains depends on its size, shape and thickness etc. Usually, quality estimation is done by human observing the grain. However, quality estimation by human is subjective in nature as it varies from man to man and quite often it is error prone. In order to standardize the quality estimation, machine-based quality estimation would be very effective. In [5] a flatbed document scanner was used to develop a computer vision based to lentil grading system. It tries to determine the lentil quality grade based only on color information. In [11], a method for grading the quality of lentil seed grains based on size uniformity is suggested. Using only color or size as a parameter may not sufficient to grade the quality of different types and varieties of lentil grains In [18], various parameters used for grading different types of lentils are given in which size is also considered as one of the important parameters.The importance of size in quality estimation of lentil is also mentioned.In this line, we intend to develop an image processing-based machine for quality estimation of the lentil grains. The machine will mainly have a hopper, a rolling belt and a camera. The grains will be fed thorough the hopper which will be spread on the rolling belt with a spreading mechanism. The spread grains will be captured as image by the camera fitted above the rolling belt. The captured image will then be processed for quality estimation. There are many parameters to be estimated for quality estimation of the lentil grains. Our basic requirement is to count the number of individual grains and determine their shapes and sizes. Before counting we need to identify the individual grains in the image. However,many grains touched or connectedbecause of imperfection in spreading mechanism,which makes it difficult to count the individual grains and determine their sizes. So, the task before us is to detect the isolated and connected grains before they are counting. The isolated grains can be counted directly but the connected grains cannot be counted directly. For counting the connected grains, the usual method is to segment the connected grains and then count them individually. There are many image segmentation methods [4,8,9,10] but most of them are very complex. In addition, they do not give good segmentation result. In this paper, we are proposing a new method which can be easily used to count the connected grains without actually segmenting them. These methods can also be used to segment the individual grains if required for size and shape determination. The method is based on the properties of convexity of an object outline. If two or more convex objects overlap and find the convex hull of the overlapped region, there are regions of convex deficiencies. The number of convex deficiencies of the overlapped region is proportional to the number of convex objects in the region. Also, convex deficiencies give appropriate place or marking for segmenting the connected objects. In [6] convex deficiency has been used to segment the connected rice grains. This property of convex hull of an object is used to identify the shape of an object [2, 7]. In our approach, we use the convex deficiencies in the convex hull of a connected component in a binary image for two purposes - to detect whether the connected component is formed by touching two or more convex objects and to count the number of connected grains without actually segmenting them. For detecting whether a connect component is formed by joining two or more grains, we find the number of convex deficiencies. If there is no or one convex deficiency, the connected component is an isolated grain. If there are two or more convex deficiencies, then the connected component is formed by joining two or more grains. The interesting property is that if the convex hull touches all connected grains, then the number of grains in the connected component is equal to the number convex deficiencies. So, by simply counting convex deficiencies, we can count the number of grains in a connected component without actually segment the grain. Another advantage of finding regions of convex deficiencies is that they give the segmentation boundaries. So, once the connected regions are identified, each individual grain can be separated if required for size determination. We have applied the method to all the captured images of lentil grains and it is giving almost perfect results except in some images in which there is one or two grains embedded inside the convex hull without directly touching the border of the hull. However, this problem could be easily solved applying some post processing techniques.

II. Generation of Binary Images

FirstBinary images are one of the most popularly used image types in various image processing applications. The main reason of using binary images is that processing binary image is much simpler than processing gray or color images. However, main issue is we do not get binary images from image capturing devices such as digital camera. Binary images need to be generated either from gray scale or color images. Generating good quality binary image itself a difficult task. Usually, binary images are generated from gray image. If the image is in color image format, it is converted into gray image before applying any image binarization method. The process of generation binary image is a two-class classification problem in which one set of pixels of the gray scale belongs to foreground class and the other set of pixels to background class. This classification of pixels of gray scale images is usually done using a threshold – those pixels less than a threshold are considered as background pixels and those greater than equal to the threshold as foreground pixels. Once the pixels of an image are classified as background and background, then binary image is generated simply by putting all background pixels to 0’s and all foreground pixels to 1’s. The question here is how to choose the appropriate value for the threshold to perform the binarization. Choosing an appropriate value for threshold is not an easy task, several researchers proposed many methods. Two of the popular methods for threshold selections are Ostu’s method [1] and Souvola’s method [3]. Otsu’s method is based on global thresholding method in which a single threshold is used to binarize a given intensity or gray image. The main problem of global thresholding is that it works well for images having uniform distribution of pixel intensity in the whole image. It does not work very well in images having variations in intensity levels in different regions of an image. Souvola’s method is based on local thresholding method to take into account the local variations in intensity values in different regions. It virtually divides a given image into different regions which are then binarized using different thresholds. It gives much better binary images when applied to old documents which have non-uniform intensity values due to aging. However, the method is more complex and the generated binary images are quite noisy and noise suppression mechanism is also applied in the method. Another threshold selection method having similar performance as that of Otsu’s is the kmeans clustering based method [16], which is simpler from the implementation point of view. These methods are based on the intensity variation of gray scale images. These methods work well when the pixels in the image have uniform distribution in intensity. However, in most of the cases, the generated binary images are quite noisy and hence usually the generated binary images need to be denoised to eliminate or reduce the unwanted noises introduced due to binarization process. Moreover, finding appropriate threshold is binarization is quite difficult..

A. Color Based Binarization

The Color objects in an image can be conveniently separated from its background using the color information rather than the intensity information. Using color information, we can selectively separate the objects of specific color or range of colors from the rest of the objects in an image. In other words, binary images corresponding to specific color objects can be generated in which the foreground regions in the generated binary image correspond to the color objects of interest and background region corresponds to the rest of the objects in a given image. However, there is one issue, that is finding suitable color model for generating good binary image. The color image in RGB format we obtain when an image is captured using a digital camera, is not suitable to generate good binary. This is because, in true sense RGB is an intensity image in the sense that all the three channels (i.e., red, green and blue) does not carry any color information. So, before applying for binarization to separate objects of specific color, the color image should to be converted into a suitable color model. Commonly used color models for binarization that have separate channels for color and intensity are YCbCr, HSV. In YCbCr color model, Y component represents the intensity, Cb and Cr are the color components. In most of the images, the Cr component is enough to separate the color objects in an image which has more red color component. In [14] fire region is separated from the other regions using only the Cr component to detect forest fire from an image in a video frame. The fire image in RGB is converted into YCbCr image, the Cr component of which mainly contains fire regions separating the greenish trees and plants in surrounding the fire. This makes it possible to exactly detect the fire region in a much simpler and faster way which is very important for a real time fire detection system. Another application of YCbCrcolor model is the detection of face regions in developing a face recognition system. In [17], skin colors of face are separated from face images to detect the face regions using the Cr chrominance component of YCbCr. . Other color model commonly used in color processing is the HSV model in which Hue and saturation caries color information and values carries the intensity information. However, this color model is not effective enough to separate color object of specific color especially the reddish and bluish color objects. In other words, YCbCr can represent the color information well as compared to the HSV color which makesthe separation of color objects from the image easier. Once a separate color components are obtained separation of the specific color object from the rest of the image can be easily done using binairzation. As compared to intensity based binarization, color based binarization is more noise free and able to separate objects of the specific color from the rest of the objects present in the input image. Moreover, threshold selection is much simpler as in most cases the histogram is bimodal in nature corresponding to color and non-color regions. The middle value of the bimodal histogram can be used as appropriate threshold to generate binary image from the Cr component. Another interesting method for generating binary image corresponding to a specific color is the using the variance information of the RGB image [15]. It is based on the concept that the variance of color pixels can carry color information.

The lentil grains available to us for ware whole lentil grains which are dark reddish in color. As the whole lentil grains have reddish color, we are using Cr component of the YCbCr color model to generate binary image to separate the lentil grains from the rest of the image. One sample of the captured RGB image of lentil grains spread on blue rolling belt is shown in Figure-1. Figure-2, shows the Cr component of the RGB image shown in Figure-1. It could be seen that in the red chrominance image, the lentil grains are have much higher intensity values as compared to the other regions of the image. This makes the separation of the lentil grains easier. The histogram of the Cr component is shown in Figure-3. The histogram is a bimodal histogram, one peak occurs at around 100 and the other pick around at 150. The middle value of between these peak values, i.e., 125 can be chosen as appropriate threshold value. The generated binary image is shown in Figure-4. It could be seen that the generated binary image is of good quality with white regions correspond to the reddish lentil grains and the black region corresponds to background region in the original input image. It may also be observed that the binary image is free from any kind of grainular noises usually introduced as a result of binaization of an intensity image. As a result no extra filtering or noise reduction method is required in the proposed method.

Alternatively, we can also effectively segment the lentil grains from the background using the Cb component of YCbCrcolor model. This is because the background color is blue which will be prominently represented in the Cb component. Separating blue background from the lentil grain image will be equivalent from the separating the grains. In other words, for lentil images having blue background and reddish lentil grains, either blue chrominance component Cb or the red component Cr can be used to separate lentil grains.

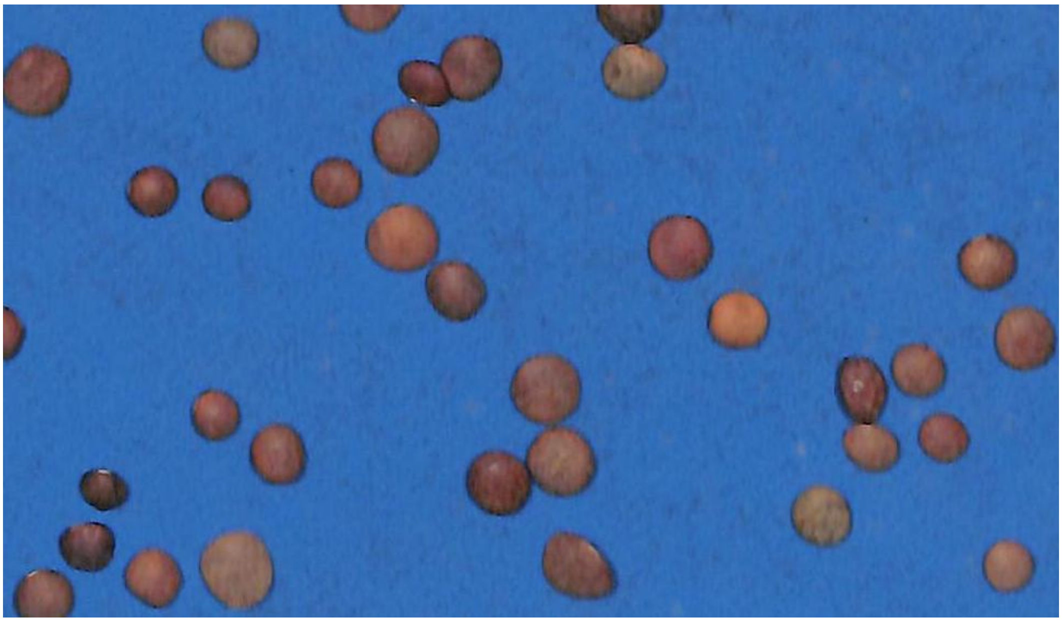

Figure-1: Lentil grains on conveyor belt

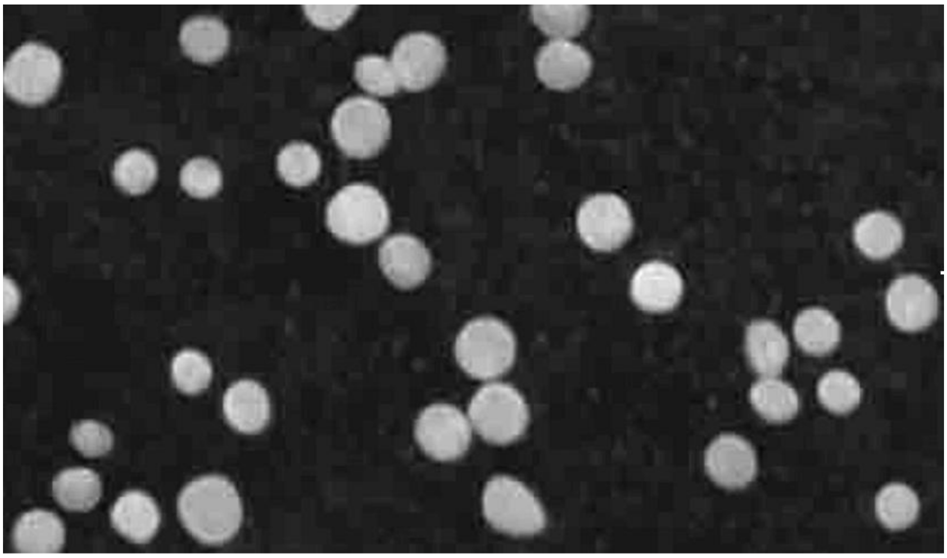

Figure-2: Cr Component of YCbCr

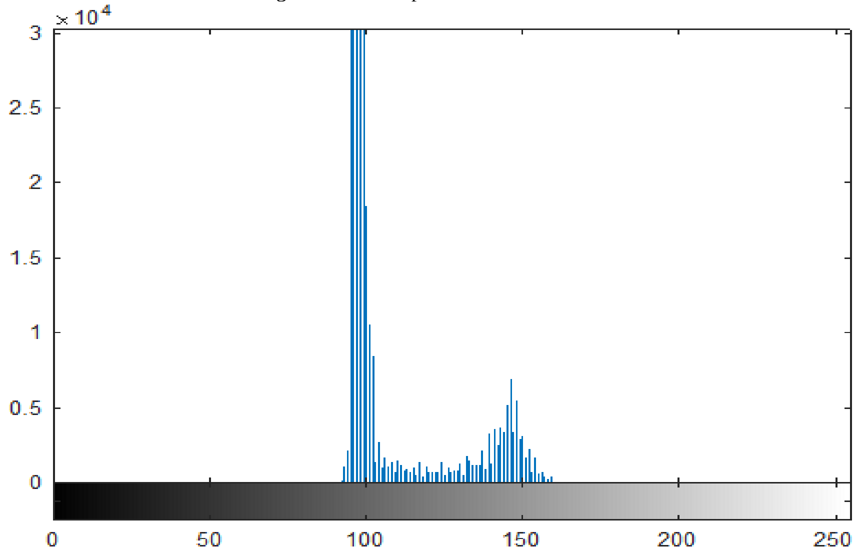

Figure-3: Histogram of the Cr Component

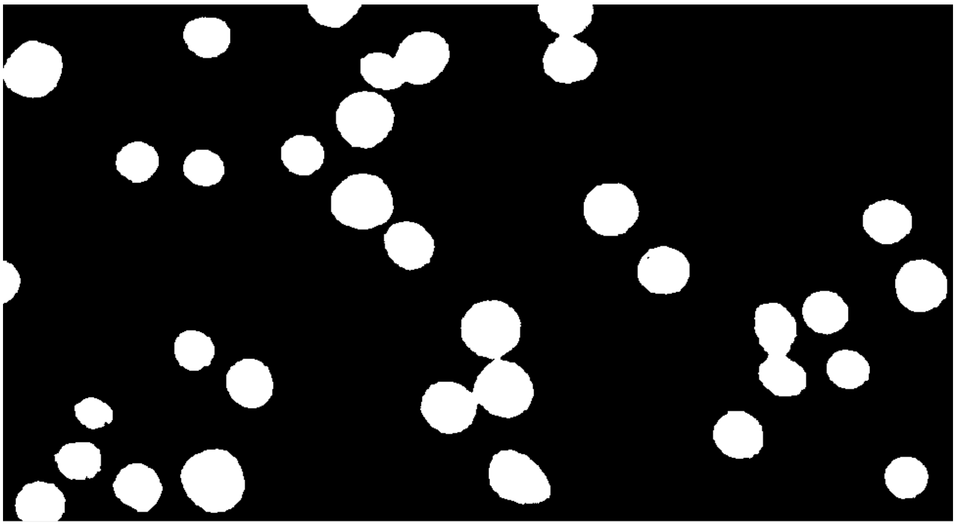

Figure-4: Binary image generated from the Cr component

III. DETECTION AND COUNTING OF CONNECTED GRAINS

We know that the binary image generated from the Cr component of the input color image has white reasons corresponding to the lentil grains. Counting these white regions in the binary image is similar to counting the grains in the original image. It may be observed that some of the white regions are formed by connection of two or grains. These connected regions will be considered as one connected component when 8-connectedness of pixel neighborhood is used for labelling the foreground objects in the binary image. So, the number of connected components will not give the number of grains. We need to find some suitable methods to correctly count the number of grains in a connected component. The usually followed method is the segmentation of the connected grains. Before applying segmentation of a connected component, we need to determine whether a connected component requires segmentation or not. Several methods are suggested to segment the separated grains in [4, 8,9,10] but these methods are quite complex and the performance is in question. As the lentil grains are rigid and have convex outlines, the most appropriate method would be to use the information about the convex deficiency of a connected component. If an outline of an object shape is not convex, the convex hull of the outline includes regions which are not the part of the object outline. These extra regions included in the convex hull of an object outline are known as convex deficiencies. If two or more convex objects are joined, the combined outline is no longer a convex and will have convex deficiencies. These convex deficiencies are useful in describing or identifying object shapes. An isolated white region corresponding to a lentil grain will have no or insignificant region of convex deficiencies. Two or more connected grains will have two or more significant regions of convex deficiencies. Using this notion, we can easily identify or detect whether connected component in a binary image is an isolated grain or formed by joining two or more grains.

Figure-5(a) shows a connected component formed by joining of two grains. The corresponding convex deficiencies are shown in Figure-5(b). It may be noted that there are two regions of convex deficiencies, white regions in the figure. Similarly, Figure-6(a) shows a connected component formed by joining of three grains and its convex deficiencies are shown in Figure-6(b). From Figures, 5(b) and 6(b), it could be seen that the number of connected grains is the same is number of convex deficiencies. This is true for all connected components formed by adjoining grains when the convex hull of the connected component touches all the grains in it. This property can be used to count the number of grains in a connected component without actually segmenting the individual grains.

Figure-5(a): Connected component of two grains

Figure-5(b): Convex deficiencies of Figure-5(a)

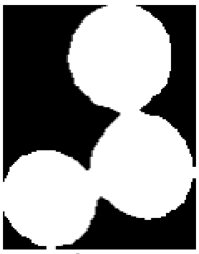

Figure-6(a): Connected components of three grains

Figure-6(b): Convex deficiencies of Figure-6(a)

A. Algorithm For Counting Grains

STEPS

1. Read the image of the captured soyabean seeds
2. Convert it into YCbCr image
3. Separate the Cr component from the YCbCr image
4. Binarize the image from the Cr Component
5. Remove the noisy smaller components (less than 300 pixels).
6. Find the connected components in the denoised binary image corresponding to the seed grains
7. Set K=0 as counter variable
8. For each connected component, compute the convex deficiencies
    a. Remove the smaller components less than 10 pixels which corresponds to border imperfection
    b. Find M, the number of convex deficiencies for each connected component

c. If $M \leq 1$, which corresponds to individual seed grains
   i. Find the individual grain in the original image corresponding to the connected component
   ii. Increment K by 1 and mark it on the individual grain
d. Else,which corresponds to the connected seed grains
   i. Find the tip points of the convex deficient points
   ii. Find the pairs the successive nearest tip points computed.
   iii. Separate the individual seed grains by the nearest tip point pairs.
   iv. Find the individual grain in the original image corresponding to the connected component
   v. Increment K by 1 and mark it on the individual grain

In the above algorithm, we read the captured grain image in RGB in Step-1. The read input image is converted into YCbCr color model from which only the red chrominance part is chosen for binarization in Steps 2 and 3. Then histogram of the Cr component is computed and appropriate threshold value is found out as explained in section –II to get binary image in step-4. The binary image in step-4 will have white regions as foreground objects i.e., grains and black region as background region. There may be smaller white components in the generated binary image which are not the parts of grains. It is found that these stray white regions have areas less than 300 pixels. So, they are removed in Step-5. The white regions are then labelled using 8-connectedness of pixel neighborhood in step-6. Once all the grain regions are labelled, each of the labelled components is accessed and processed for counting in step-6. We then set a counter variable K to 0 in step-7. In step-8, we find each connected component, compute the convex hull and then convex deficiencies. When finding convex deficiencies from the convex hull, some small traces of white pixels are introduced in case of isolated grains because some grain shapes are not perfectly round or convex. To avoid such small traces of white pixels, connected white regions less than 10 pixels are removed in step8(a). This removes small white regions which are not the part of any convex deficiencies. Then, the number of convex deficiencies M is computed in step-8(b) to find the number numbers of grains in the connected component. In this step, the convex hull of the connected component is computed first which is xored with the connected components to find the convex deficiencies. The number of convex deficient regions can easily determine whether the connected component is an individual seed grain or connection of two or more seed grains. For individual seed grains, there is no convex deficient regions. For connected grains, the number of convex deficient regions are two or more. So, in step-8(c), we test whether the connected component is individual grain, if so the location of the grain in the original image is identified and then, the counter is incremented by 1 and then the counter value is marked on the identified grain. If the connected component consists of two or more seed grains in step-8(d), the tip points of the convex deficient regions are found out. Then find the pairs of shortest tip points. Each pair corresponds to separating line to segment the connected grains in the connected component. Once a grain is segmented, its location is identified and the incremented value of the counter variable is marked on the identified grain. In this way, we can count the number of seed lentil grains in a captured image.

IV. EXPERIMENTAL RESULTS

For the experimental purpose, we consider the 16 different images of lentil grains captured from the conveyer belt on which lentil grains are spread mechanically. Each image is RGB color image having dimension of 3506 x 4961 x3 which is very big for display on a computer screen or a page. The image needs some pre-processing as it has border areas of the conveyer belt which do not contain grains. So, the grain regions are cropped from the original image to use as input for further processing.The cropped image is then converted into YCbCr color model to separate the color components in the image. The color components consist of blue chrominance (Cb) and red chrominance (Cr). For the images considered, either of the chrominance components can be used to separate the foreground objects from the background region. However, we use Cr component to separate the reddish lentil grains from the background region in the form of a binary image. After separating the lentil grains from the background, grain components are extracted using connected component analysis. Each of the connected component is further processed to detect whether a connected component is an isolated grain or a bunch of connected grains. This is done computing the number of convex deficiencies of the connected component. First the the convex hull is computed and convex deficiencies are computed by xoring the convex hull with the connected component. The number of convex deficiencies is then used to detect whether the connected component is an isolated grain or a connection of two or more grains. There are various methods of computing convex hull of points in 2-dimensional planes as described in [12, 13]. MATLAB has a built-in function named bwconvhull, which computes the convex hull of an input binary image and returns the convex polygon region as binary output. By xoring the binary image of the connected component with from the binary convex hull using bitxor function in MATLAB, we get the binary image corresponding to the convex deficiencies. We the use gain the connected component analysis to count the number of convex deficient regions. If there is no convex deficient regions, the connected component is an isolated grain. If there are two or more convex deficient regions, the connected component consists of two or more grains. For connected grains, individual grains are isolated by separating them at the nearest successive pairs of tip points of the convex deficient regions. A counter variable is used to count the number of grains in the input image. The proposed algorithm has been implemented in MATLAB and tested on all images considered. It has been found it can perfectly detect all the connected grains regardless of its size. Also, it can exactly count the number of grains in most of the images. For display purpose, we show here two images as shown in Figures 7(a) and 8(a). These images are cropped images from the original images of size 3506 x 4961 x3 taking mainly the portion containing connected seed grains which are then resized to fit in the page. The cropped image in Figure-7 is smaller and hence its size less reduced to fit the page. So, the seed gains in

Figure-7(a) looks a little bigger as compared to seed grains in Figure-8(a). The counted seed grains are respectively shown in Figure-7(b) and Figure-8(b). It may be seen that all the seed grains are counted correctly in both images. The counting is done in the left to right fashion. That is the leftmost grain is counted as first grain and the rightmost grain is the last grain. From Figure-7(b), we see that each seed grain is correctly counted by appropriately marking them and the number of seed grain in the image is 45. Similarly, the number of seed grains in correctly counted as 56.

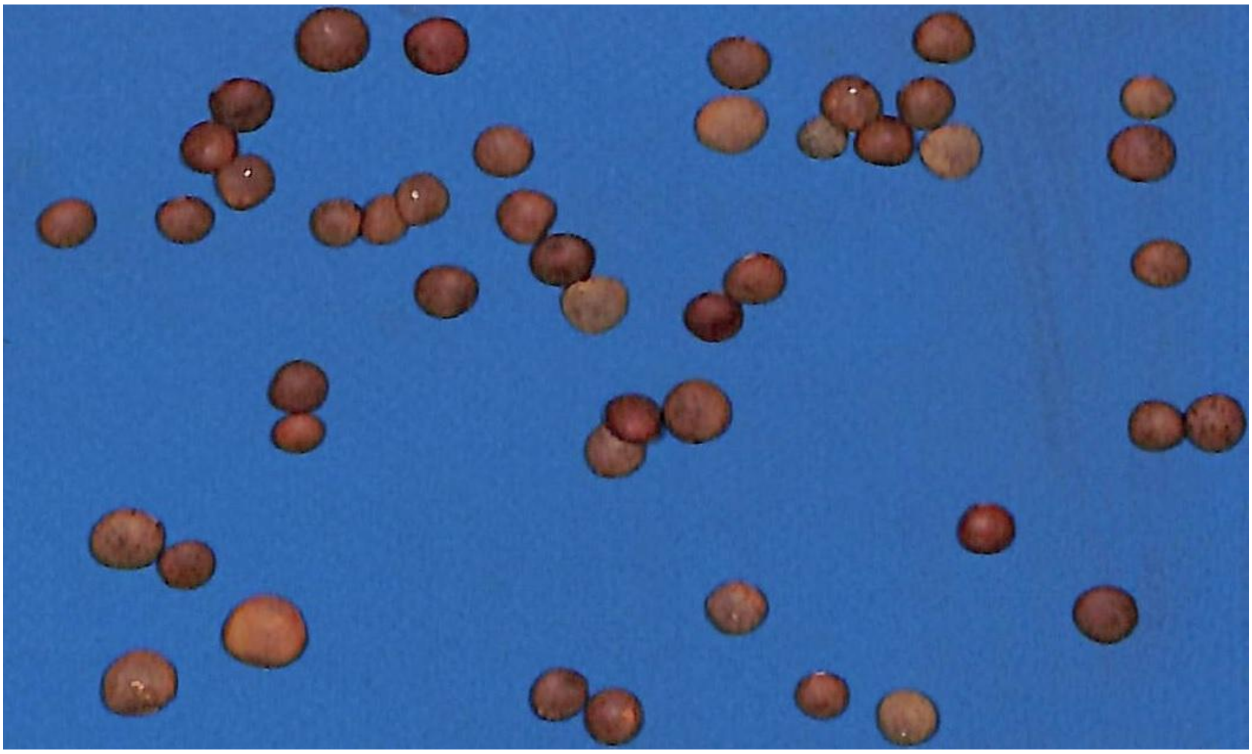

Figure-7(a): Image-1 of Lentil grains captured from the conveyor belt having blue color.

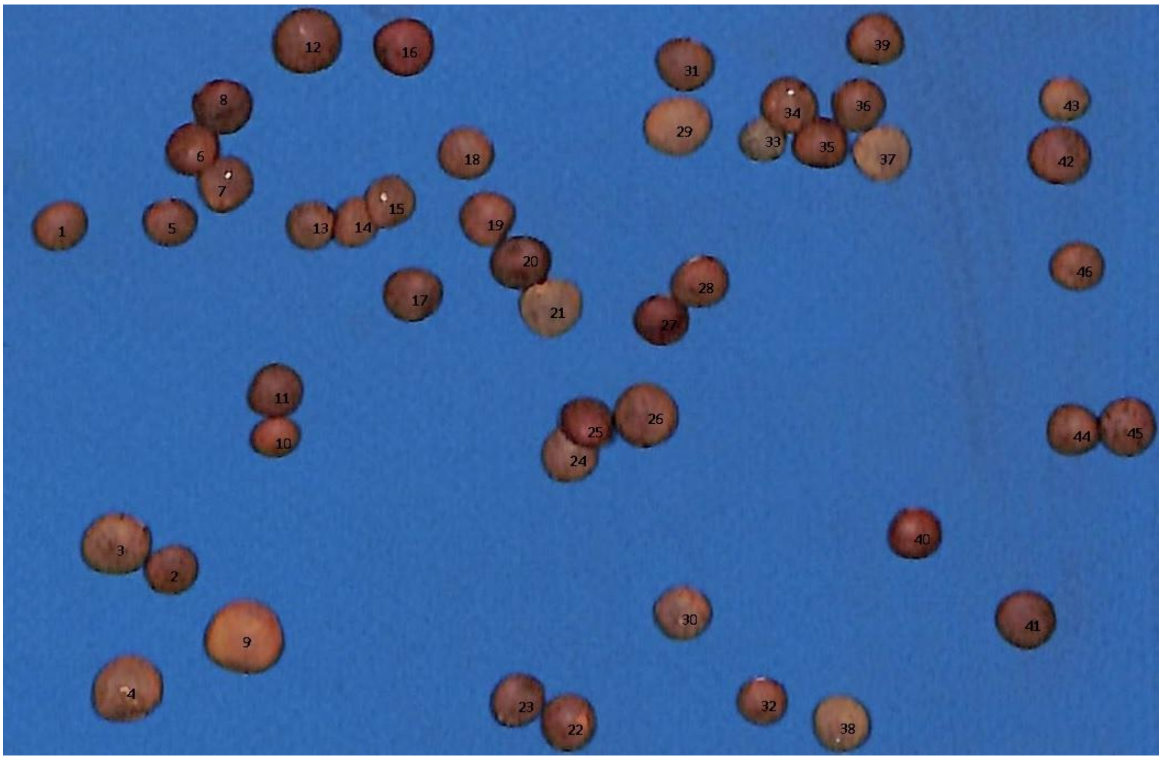

Figure-7(b): Counted lentil grains (45) of Image-1 in Figure-7(a).

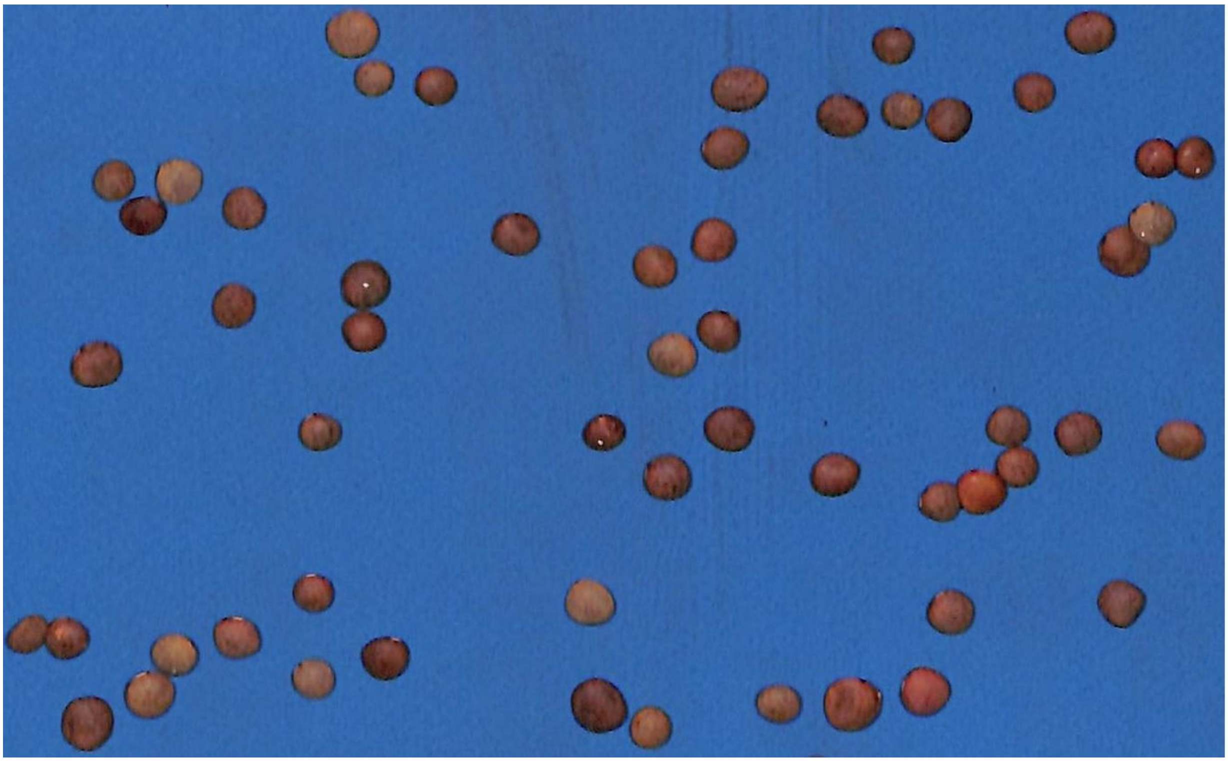

Figure-8(a): Image-2 of captured lentil grains on conveyor belt having blue color.

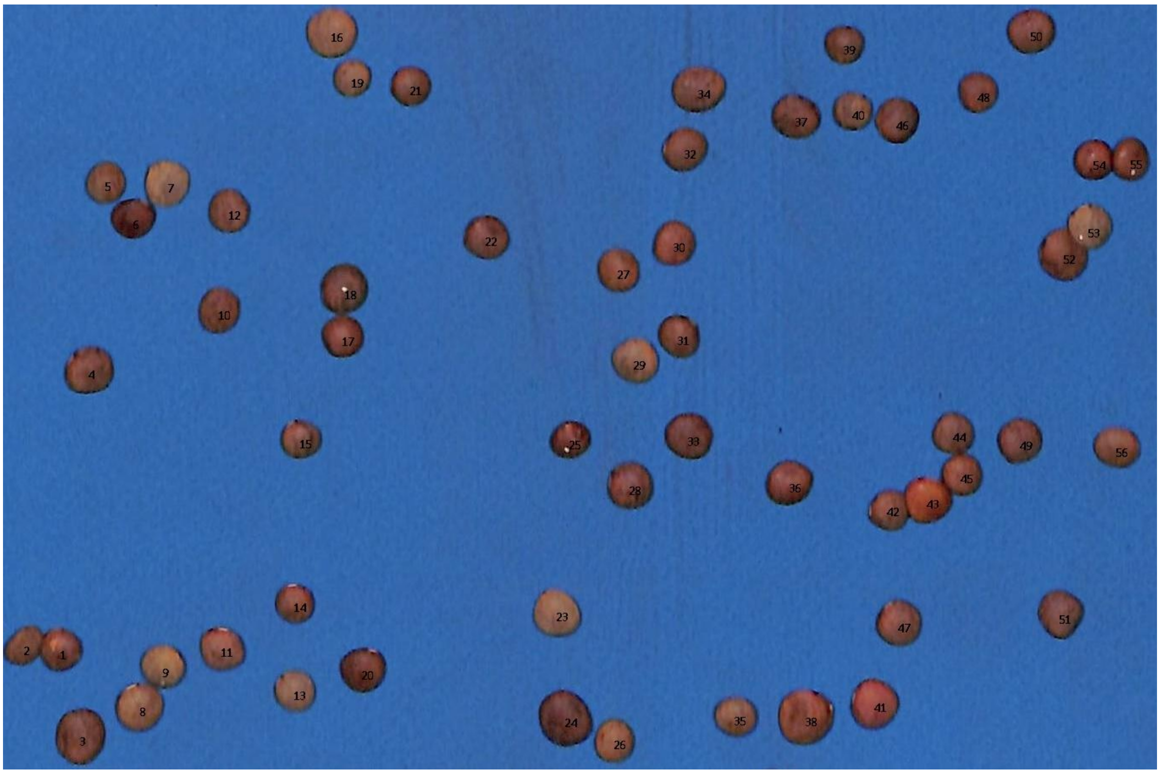

Figure-8(b): Counted lentil grains(56) in Image-2 of Figure-8(a).

V. CONCLUSIONS

A simple and fast convex deficiency-based method for detection and counting of lentil grains has been proposed in the paper. The proposed method is based on generating good quality binary image from the red chrominance of the YCbCr color model. The grains are then separated from the background using the connected component analysis. Each separated connected component is then checked whether the component is formed by isolated seed grains or joining of two or more grains by finding the convex deficiencies. The nice part of the proposed method is that by joining the nearest successive pair of tip points of convex deficient regions in the convex hull of a connected component, we can separate the grains enclosed in the convex hull. It has been tested in all the captured lentil grain images and all connected are detected as expected. So, far it can correctly count the total number of grains in all captured images. In fact, all captured images are of good quality seed grains, we need to do more experimentation, whether the proposed algorithm will correctly count all grains in images having poor quality seed grains as well. In general, the proposed method can be successfully used to count the other types of grains having convex shapes.

REFERENCES

[1] Otsu, N.: A threshold selection method from gray-level histogram. IEEE Transactions on Systems, Man and Cybernetics, vol 9(1), pp. 62–66 (1979)

[2] Parker, J.R.: The use of convex deficiencies for the recognition of hand printed digits. In: Melter, R.A., Wu, A.Y. (eds.) Vision Geometry III, vol. 2356, pp. 169–175, 1994

[3] Sauvola, J., Seppanen, T., Haapakoski, S., Pietikainen, M.: Adaptive document binarization. In: Fourth International Conference Document Analysis and Recognition (ICDAR), Ulm, Germany (1997)

[4] N. S. Shashidhar, D. S. Jayas, T. G. Crowe and N. R. Bulley, "Processing of digital images of touching kernels by ellipse fitting," Canadian Agricultural Engineering, vol. 39, no. 2, pp.139-142, 1997.

[5] Shahin M.A., Symons S.J. "A machine vision system for grading lentils. Can. Biosyst. Eng"., vol. 43, pp. 7–14, 2001.

[6] S. Q. Yang and D. J. He, "Automated identification and separation of touching rice grains with machine vision," Journal of Agricultural Mechanization Research, no. 3, pp. 62-65, 2005.

[7] Arcelli, C., Sanniti, G., Svensson, S., " Computing and analysing convex deficiencies to characterise3d complex objects", Image Vis. Comput. 23(2), 203–211, 2005.

[8] Y. Ling, Y. M. Wang, M. Sun and X. C. Zhang, "Application of watershed algorithm to paddy image segmentation," Transactions of the Chinese Society of Agricultural Machinery, vol. 36, no. 3, pp.95-98, 2005.

[9] Y. B. Hu and J. P. Liu, "Cell image segmentation of gastric cancer based on region – growing and watersheds," Computer and Digital Engineering, vol. 34, no. 8, pp. 151-154, 2006.

[10] Qing Yao, Yingfeng Zhou, Jiangning Wang, "An Automatic Segmentation Algorithm for Touching Rice Grains Images", International Conference on Audio, Language and Image Processing (ICALIP2010), pp. 802-805, 2010.

[11] L.S. LeMasurier, J.F. Panozzo, C.K. Walker, " A digital image analysis method for assessment of lentil size traits", Journal of Food Engineering, vol 128, pp-72-78, 2014..

[12] M. A. Jayaram, Hasan Fleyeh, Convex Hulls in Image Processing: A Scoping Review, *American Journal of Intelligent Systems*, Vol. 6 No. 2, 2016, pp. 48-58. doi: 10.5923/j.ajis.20160602.03.

[13] Gamby, A.N., Katajainen, J. (2019). A Faster Convex-Hull Algorithm via Bucketing. In: Kotsireas, I., Pardalos, P., Parsopoulos, K., Souravlias, D., Tsokas, A. (eds) Analysis of Experimental Algorithms. SEA 2019. Lecture Notes in Computer Science(), vol 11544. Springer, Cham. https://doi.org/10.1007/978-3-030-34029-2_30

[14] Y. Kirani Singh and Debasish Deb, "Detection of Fire Regions from a Video Image Frames in YCbCr Color Model", International Journal of Recent Technology and Engineering (IJRTE), Volume 8, No. 8, September, 2019.

[15] Yumnam Kirani Singh, "Detection of Fire Regions Using RGB color Variance", International Journal of Research in Engineering and Science (IJRES) ISSN (Online): 2320-9364, ISSN (Print): 2320-9356, www.ijres.org Volume 9 Issue 6, PP. 10-14, 2021.

[16] Yumnam Kirani Singh, "A Simple and Effective method for Segmenting Lung Regions from CT Scan Images using K-Means", Presented to International Conference BigBML2021, NIT Silchar, 19-20, December 2021.

[17] VanlalHruaia, YumnamKirani Singh, "Detection of Face detection from only the Chrominance component", SSRN Electronic Journal, February 2022, Doi:10.2139/ssrn.4100218

[18] Marco Del Coco, Barbara Laddomada, et. al., "Characterization of a Collection of Colored Lentil Genetic Resources Using a Novel Computer Vision Approach", https:// https://www.mdpi.com/journal/foods, https://doi.org/10.3390/foods11243964 pages- 15, 2022